

EXHIBIT 2

Faculty Directory

[Academic Schools and Departments](#)

[Research Centers](#)

Tomek Strzalkowski



Professor, Cognitive Science

Prof. Tomek Strzalkowski research interests span a wide spectrum of human language technology including computational linguistics and sociolinguistics, socio-behavioral computing, interactive information retrieval, question-answering, human-computer dialogue, serious games, social media analytics, formal semantics, and reversible grammars. He has directed research sponsored by IARPA, DARPA, ARL, AFRL, NSF, the European Commission, NSERC, as well as a number of industry-funded projects. He was involved in IBM's Jeopardy! Challenge in advanced question answering. Dr. Strzalkowski has published over a hundred and fifty scientific papers, and is the editor of several books, including *Advances in Open Domain Question Answering*. He serves on the Editorial Board of the journal of *Natural Language Engineering*.

Prior to joining RPI, Dr. Strzalkowski was Professor of Computer Science at SUNY Albany. At SUNY, he was the founding Director of the Institute for Informatics, Logics, and Security Studies with research budget of more than \$35 million. He came to SUNY from GE CRD where he was a Natural Language Group Leader and Principal Scientist. At GE, Dr. Strzalkowski directed projects on automated technical manuals, medical informatics, speech recognition, automated summarization, as well as multimedia processing including language and video. Before coming to GE, he was a research faculty at the Courant Institute of New York University, where he worked on applications of natural language processing to information retrieval.

Current projects include research into social dimensions of information spread online, internet ethnography, and building effective AI defenses against disinformation and exploitation of human socio-cognitive vulnerabilities online, including social engineering attacks. Some example projects include:

- GATOR: The Goal-oriented Autonomous Dialogue System. We develop a new type of human-machine dialogue system that uses deep learning technologies (such as transformers) to learn how to recognize and generate dialogue plans, i.e., semantic and pragmatic structures that represent one party's goals and intentions, as well as the impact these are having on the other party. Unlike the current transformer-driven chatbots, the core learning is not to transform one language expression (input) into another language expression (response) but instead to construct a response plan that would properly address the plan in the input and the history of interaction. Consequently, the learning process takes three types of information: (1) the input utterance; (2) its semantic-pragmatic plan, i.e., the plan that was used to produce the utterance, and (3) the history of interaction up to this point. Furthermore, the cumulative history of the dialogue is not merely the memory of the utterances exchanged earlier, but

it captures, in a condensed semantic form, the evolving state of the parties’ objectives as well as the emerging sociolinguistic behavioral patterns of both (all) parties.

- Personalized AutoNomous Agents Countering Social Engineering Attacks (PANACEA) protects online users against current and future forms of social engineering. PANACEA serves as an intermediary between attackers (human, automated, hybrid, coordinated) and the potential victim(s) they target. Depending upon the nature and source of communication, PANACEA either handles it autonomously, or allows the user to proceed with an exchange while monitoring the conversation and intervening as needed by (1) inserting or modifying users’ messages, (2) instructing the user how best to respond, while at the same time (3) initiating an investigation to identify the attacker. (DARPA ASSED Program)
- COMETH (Computational Ethnography from Metaphors and Polarized Language). The objective of this project is to develop a methodology and accompanying software tools for constructing dynamic socio-behavioral models of communities based on online content that their members produce. A community can be defined by the set of salient concepts that its members recognize, along with the values they assign to them. The resulting causal models are then applied to derive culturally biased interpretations of novel information by prototyping the process by which such new information is adapted to fit into the community current model. (DARPA UGB)
- Social Convo: A New Approach to Modeling Information Diffusion in Social Media. In this project, we recast our understanding of all social media as a landscape of collectives, or “convo”: sets of users connected by a common interest in an (possibly evolving) information artifact, such as a repository in GitHub, a subreddit in Reddit or a group of hashtags in Twitter. Convo are represented by the collections of features that capture their internal social dynamics. Furthermore, convo are basis for modeling large and small internet-based communities as “hybrid organisms” that interact in various ways with one another and react collectively to external stimuli, including information and disinformation campaigns. (DARPA SocialSim)

Education

- Ph.D. Computer Science, 1986, Simon Fraser University, Canada

Research Focus

Artificial Intelligence
Natural Language Processing
Computational Sociolinguistics

Contact Information

tomek@rpi.edu
http://lacailab.cogsci.rpi.edu/

Select Works

Rosa Mikeal Martey, Adrienne Shaw, Jennifer Stromer-Galley, Kate Kenski, Benjamin Clegg, James Folkestad, Tobi Saulnier, Tomek Strzalkowski (2017) Testing the Power of Game Lessons: The Effects of Art Style and Narrative Complexity on Reducing Cognitive Bias. International Journal of Communication, Vol. 11, pp. 1635-1660.

Tomek Strzalkowski and Carl Symborski (2017) Lessons Learned About Serious Game Design and Development. Games and Culture. Vol. 12(3) 292-298. Sage.

Tomek Strzalkowski, Teresa Harrison, Ning Sa, Gregorios Katsios, and Ellisa Khoja (2018) GitHub as a Social Network. Proceedings of 9th AHFE Conference, Orlando, FL. Springer Nature.

Tomek Strzalkowski, George Aaron Broadwell, Sarah Taylor, Laurie Feldman, Boris Yamrom, Samira Shaikh, Ting Liu, Kit Cho, Umit Boz, Ignacio Cases and Kyle Elliott. 2013. Robust Extraction of Metaphors from Novel Data. Proceedings of NAACL Workshop on Metaphors in NLP. Atlanta, GA.

Small, Sharon, Jennifer Stromer-Galley and Tomek Strzalkowski (2011) Multi-Modal Annotation of Quest Games in Second Life. Proc. of ACL-2011, Portland.

George Aaron Broadwell, Jennifer Stromer-Galley, Tomek Strzalkowski, Sarah Taylor, Umit Boz, Samira Shaikh, Liu Ting, and Nick Webb (2013) Modeling Socio-Cultural Phenomena in Discourse. Journal of Natural Language Engineering. 19(2): 213-257. Cambridge.

T. Strzalkowski & S. Harabagiu (eds.) Advances in Open-Domain Question Answering. Springer, 2006


SEARCH FOR FACULTY:

enter a name | select from the dropdown

or

SEARCH BY RESEARCH FOCUS:

Search



110 Eighth Street
Troy, NY USA 12180

SCHOOLS & PROGRAMS

Architecture

Business

Engineering

Humanities, Arts, & Social Sciences

IT & Web Science

RPI CONNECTIONS

Admissions

Alumni/ae & Friends

Athletics

Human Resources

Library

ENGAGE

Apply Now

Contact Us

Give to Rensselaer

Visit Campus

Send Feedback

